Домашнее задание №5

Выпуклость функции потерь логистической регрессии
(частный случай: в задаче бинарной классификации)

Срок сдачи: пятница, 17.04.2020
Вес задания: 0.2
Максимальный номинальный балл: 100
Максимальный итоговый балл: 20

НАПОМНИМ:

Логистическая регрессия - частный случай обобщенной линейной модели, в которой параметр $p(\theta, x_i)$ распределения Бернулли, инетрпретируемый как вероятность объекта x_i быть помеченным как класс "1", вычисляется следующим образом:

$$p(\theta, x_i) = \frac{1}{1 + \exp(-\theta \cdot x_i)}$$

В соглашении, что классы кодируются как $y_i = 0$ для "отрицательных" примеров и $y_i = 1$ для "положительных", правдоподобие выборки $\mathcal{T} = \{y_i, x_i\}$, $i = 1...N$ записывается следующим образом:

$$L(\mathcal{T}, \theta) = \prod_{i=1}^{N} \left(p_i^{y_i} * (1 - p_i)^{(1-y_i)} \right)$$

Тогда логарифм правдоподобия записывается следующим образом:

$$\begin{aligned} \ell(\mathcal{T}, \theta) &= \sum_{i=1}^{N} \left(y_i * \ln p_i + (1 - y_i) * \ln(1 - p_i) \right) = \\ &= -\sum_i \ln\left(1 + \exp\left(\theta \cdot x_i\right)\right) + \sum_i y_i * \theta \cdot x_i \end{aligned}$$

Вспомним, что согласно принципу максимизации правдоподобия (maximum likelihood, ML), функция потерь в этой задаче совпадает по модулю с логарифмом правдоподобия $\ell(\mathcal{T}, \theta)$, но противоположна с ней по знаку:

$$\mathcal{L}(\mathcal{T}, \theta) = -\ell(\mathcal{T}, \theta)$$

Дополнительные сведения о логистической регрессии в задаче бинарной классификации можно почерпнуть в материалах к Лекции №11 или в рекомендуемой литературе.

ЗАДАНИЕ:

Показать, что функция потерь $\mathcal{L}(\mathcal{T}, \theta)$ - выпуклая на всем пространстве Θ